

DOCUMENT RESUME

ED 268 182

TM 860 249

AUTHOR Rudman, Herbert C.
TITLE Testing Beyond Minimums. Occational Paper #5.
INSTITUTION Association of State Assessment Programs.
PUB DATE 85
NOTE 38p.; Paper presented at the Michigan School Testing Conference (25th, Ann Arbor, Michigan, February 26-27, 1985) and at the Education Commission of the States/Colorado Department of Education Assessment and Policy Conference (Boulder, Colorado, June 10-13, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142) -- Collected Works - Serials (022)
JOURNAL CIT ASAP Notes; spec iss Spr 1985
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Standards; Basic Skills; *Cognitive Processes; *Educational Assessment; *Educational Testing; Elementary Secondary Education; *Minimum Competencies; Minimum Competency Testing; Problem Solving; *Public Policy; *State Programs; Testing Problems; Testing Programs
IDENTIFIERS Excellence in Education; *Mandated Tests; *Test Score Decline

ABSTRACT

This paper considers three areas related to the testing of minimum knowledge and skills: public policy, the nature of higher-order thinking and selected aspects of educational measurement. The discussion of public policy issues centers around an apparent dichotomy of egalitarianism and elitism in public education, schooling and the public economy, and the pendulum affect on policy and development. The topic of higher-order thinking is limited to the hierarchical nature of learning, the impact of coaching on this hierarchy, and issue in the instruction of problem-solving techniques. Two aspects of educational measurement are described: (1) the nature of test score decline and (2) the impact of mandated assessment on achievement. It is concluded that while schools are not powerful enough to affect the change of raising levels of academic excellence beyond the minimum on a national level, a powerful tool for accomplishing this has been statewide assessment programs. In regard to the State of Michigan's educational program, state action to improve educational standards should include the following steps: (1) develop a plan which incorporates a two-tier evaluation of student academic status; (2) establish a standard-setting advisory committee; (3) assume a major portion of funding for mandated assessments; and (4) establish a technical advisory committee. A reference list is appended. (PN)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ASAP NOTES

Spring
1985

Association of State Assessment Programs Newsletter

Editors: C. Thomas Kerins
Mervin M. Brennan

Associate Editors:

Leslie J. Fyans, Jr.

Kenneth R. Redding

Norman K. Stenzel

Carmen Woods Chapman

Contributing Editors:

Rose M. Owens

Edward D. Roeber

SPECIAL EDITION Occasional Paper #5

Testing Beyond Minimums by

Herbert C. Rudman
Michigan State University

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

H. C. Rudman

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

Reprinted with permission

ED268182

TM 860 249

TESTING BEYOND MINIMUMS ^{1.}

Herbert C. Rudman
Michigan State University

If any of us are still under the delusion that state-mandated assessment activities have little influence on educational programs, this session may end that notion. This is not to say that the power of state-mandated assessments is an absolute power to do good or mischief, but it most certainly does affect teaching and learning for the long term. How we harness that power will, ultimately, determine the future of education not only in this state, but in those other states where testing has been used to affect the nature of the learning experiences of students and the instructional goals of their teachers.

For more than twenty years now we have been aware of the phenomenon of falling test scores across a variety of academic subjects (Harnischfeger and Wiley, 1975; Munday, 1979; Lerner, 1979; Tomlinson and Treacy, 1979; Wainer, 1984). The phenomenon of falling test scores is not new, nor is it limited to students within the United States alone (Goldstein, 1983; Gipps and Goldstein, 1983; Start and Wells, 1972). During the 1970's, in particular, considerable debate raged in Great Britain over the apparent decline in educational standards as evidenced by falling test scores as measured by the Assessment of Performance Unit, a British counterpart to our own National Assessment of Educational Progress.

A number of good and valid reasons have been presented for the apparent deterioration of school learning (Shane, 1977; Wirtz, 1977), but no publication

1.

A paper presented at the 25th Michigan Testing Conference, University of Michigan, February 27, 1985.

has had the political and educational impact upon the schools as has the report of the National Commission on Excellence in Education, A Nation at Risk (1983). One measure of its influence is the recent spate of articles and speeches denouncing its conclusions and its recommendations. Another measure is the number of spin-off pieces which have emerged from such diverse sources as the Michigan Republican Party's "Action Plan for the '80s" (1983), the Report of the Michigan Commission on High Schools (1983), the Michigan State Board of Education's Better Education for Michigan Citizens: A Blueprint for Action (1984), a report of a special task force appointed by Michigan Governor James Blanchard entitled, Governor's Educational Summit Task Force (1985), and a spate of similar reports from other political and educational organizations across this country (Pipho, 1984).

It has usually been assumed by those addressing this topic that this general drop in student achievement has been the result of a deterioration among students of average and less-than-average ability. It has also been assumed that the high achieving students were continuing to score well on achievement tests (College Entrance Examination Board, 1975, 1976, 1977). These were the students for whom little concern was felt since they were believed to be self-motivating learners who took advantage of their experiences outside of as well as inside the schools. Recent analyses of test score data have been conducted across ability cohorts or groups and a different picture has emerged. It would appear that the most dramatic drops in test scores, by any measure of school subject analyzed, have occurred within the high achieving groups (College Entrance Examination Board 1982, 1983; National Assessment of Educational Progress, 1983; Eberly, 1976; Brandenburg, Gallagher and Berkheimer, 1978; Munday, 1979; Sederburg and Rudman, 1984).

Since this high ability cohort will provide the technical and professional leadership within two decades or so, we should be deeply concerned about the patterns of student achievement that have developed over the years. Recent months have witnessed increased activity and concern for the high ability student, but this concern is focused less on the drop in achievement and more on expanding the teaching and measurement of higher-order thinking. This, in turn, has generated a rekindled interest in cognitive psychology (Fredericksen, 1984; Morante and Ulesky, 1984; Beyer, 1984; de Bono, 1983; Paul, 1985; Scriven, 1985; Lipman, 1985). Through the happy confluence of educational concerns interacting with the exercise of political power, teaching, learning and testing have come to the attention of educational practitioners in a comprehensive package of focused concerns rare to those who are weary of educational pendulums and fads that too often focus on one small facet of an educational problem.

However, there is a dark side to all of this new popular interest in the measurement and teaching of higher order skills. The same old pattern of mechanistic thinking, simplistic solutions, and the emergence of "gurus" is evident as one reads the usual journals, attends the usual conferences and staff meetings. We do have a chance to address a serious educational issue. If we eschew instant solutions, are selective of the paths we follow, and are deliberative in our planning we might increase the probabilities of a long-term resolution of a serious social and educational issue. Political leaders are often impatient with deliberative study; they are driven by deadlines. Educational leaders are too used to jumping on and off bandwagons. If both can continue to reinforce each other, in this instance, we might yet emerge with a significant addition to educational practice.

It can be argued that the schools of this nation have been effective in

meeting the tasks that have been presented to them by the society at large, and by those who determine public policy at all levels of government. Since 1964, social policy has dictated that the schools address the needs of designated target populations who have suffered neglect for more than a century and a half. Much of the social legislation which has emanated from Congress and from state legislatures has focused on entitlement programs targeted at minimum social and educational problems. Children who have rarely had the preschool experiences to enable them to benefit from regular educational programs are now given opportunities to gain experiential bases from which subsequent schooling can be enriched. Public television is used to present cognitive preschool experiences and to open an unfamiliar world to children raised in poverty. Minimum educational, social and nutritional benefits are provided in the hope that children will have a better base for school learning. And for the past two decades, the schools have responded magnificently to this need. Not since the last wave of mass immigration ended in the early 1920's have the schools had such a clear mandate and not since that time have the schools been able to demonstrate so clearly how effectively they can respond. But herein lies the rub.

We have done such an effective job of addressing our minimal needs for these past twenty years, that our schools and those who staff them have, too often, made of the minimum, the maximum. This is reflected in the awarding of certificates of merit to those Michigan schools whose students attain minimal performance in state-mandated measures of achievement in reading and mathematics, in the emphasis on lower vocabulary loads in textbooks used throughout the schools of this nation, and in the pacing of instruction of classrooms at all levels of schooling from the kindergarten through the university (though not necessarily all classrooms). All this, and more, has

resulted in lowered expectations of students, and in an imbalance in the targeting of differentiated instruction among those of varying abilities.

Is it possible for a society which passionately believes in equal opportunity for all of its citizens to differentiate instruction in the schools without reverting to an intellectual elitism? The answer to this question must be an affirmative one if one truly believes in the worth of the individual; a hallmark of a democratic society.

The concern voiced by many for testing beyond the minimum is a complex one. It is not a problem easily considered nor simplistically addressed. It is a topic that has embedded within it serious questions of political philosophy, cognitive psychology, educational measurement, and instructional methodologies. There is no way that this paper can do justice to the topic, for the issues involved are not amenable to short-term analysis or discussion. What I would hope to do is broaden the discussion and suggest some alternatives we need to consider before making a long-term commitment.

This paper will consider three areas related to the more immediate topic of expanding our preoccupation with testing of minimum knowledge and skills; public policy, the nature of higher-order thinking, and selected aspects of educational measurement. Some attention will be given to the status of current knowledge about these three areas, and to some tentative suggestions as to how we might address them.

What Do We Know?

Several questions come to mind when considering the current concern with measurement of higher-order thinking. One, has to do with how these measures will be used. Will they be used only for classroom instruction that will place no greater nor less importance upon these skills than any other skills, or will they be used as monitoring instruments to reflect public concerns and policies

about "excellence" in education? If they reflect a public concern, to what must we be sensitive in policy development? A second consideration, is the definition of those clusters of higher-order skills we refer to as "creative thinking", "inquiry", or "critical thinking". And finally, what measurement problems and issues must we face if we are to develop tests that are responsive to public demands for an education that extends beyond the minimum?

Public policy issues. For the sake of simplicity my discussion of public policy issues will center around (1) an apparent dichotomy of egalitarianism and elitism in public education, (2) schooling and the public economy, and (3) the pendulum affect on policy development.

1. Egalitarianism vs. elitism. The most serious discussions of philosophy, theology or politics often reveal contradictory arguments. We face no less a set of contradictions when serious discourse begins on the differentiation of instruction and programs for students in the public schools. On the one hand, we insist on individualizing instruction whenever possible, we identify target groups for special instructional attention, and insist that all students must have an equal opportunity to learn all of which they are capable. But at the same time, we resist tracking students of different abilities — measured or real —, we reflexively assume that students held behind their classmates for special instruction will be unfairly labeled, we insist that all students have equal access to all curricula, and hold that suggestions to reallocate human and fiscal resources for high achieving students, are tantamount to the introduction of elitism in the schools.

Clark (1985) in a cross-national discussion of the linkage between high schools and the university, identified this contradiction as the source of bias against excellence in school programs.

The distinctively American problem posed by universal education springs from an excess of hope. A strong belief exists both within and outside the U.S system of education that the schools have failed if all members of an age group do not persist in school for 12 years and graduate with some kind of high school diploma. ...

...there is the price that Americans pay for their extremely high (social) expectations. The dream of universal education ... that includes all students and all curricula and that brooks no individual failures, is a huge burden on the school system. The democratic idealism implicit in the concept of the comprehensive school has contributed significantly to community and social integration. ... But the perceived failure of U.S. secondary schools is ... the result of a shortfall in outcomes when they are measured against the expectation that the comprehensive school ..can be helpful to all students from all types of backgrounds, ... while at the same time minimizing invidious distinctions.

...The more that international standards of academic achievement shape American expectations for their schools, the more it can be said that the structure of schooling in the U.S. is deeply and systematically biased against "excellence". (pp. 392, 395-396)

We have tended to equate an abiding concern for equality of educational opportunity, and an inherent desire to improve the learning of students in the low-achieving cohort, with democratic practice while at the same time condemning as "elite" those same endeavors when applied to programs for students who demonstrate high measured ability (Duncan, 1980). It is this general fear of elitism which makes me shy away from the more traditional terms we use of "gifted", and "talented". Focusing instruction on children who demonstrate a high propensity for school learning is no more an elitist concept than addressing the needs of children who need remediation. Both target populations use the same instructional methodology of prescription and diagnosis. Both targets are the recipient of a democratic philosophy that proclaims the fundamental worth of the individual. It has been no more nor less "democratic" to focus a large share of our human and fiscal resources on the minimal needs of all students than it would be to reallocate our resources so that all students

get an equitable share of instructional time and programs that are appropriate for their abilities. Reducing instruction to the lowest common denominator is not egalitarian, it is destructive of the worth of the individual.

The enigma that we face is that we deal with constants; time, money, and human resources. If all students are to have equal opportunity to succeed and to learn all of which they are capable, then these constants cannot remain "constant" for all students. Low achieving students may need more time to learn. If this means a longer stay in school, so be it. If high ability students need more focused instructional time, then that should be supplied. Both instances, of course, mean a change in the "time constant". Since time, in this instance, does require modification of a standard school day, or a longer school year, or a longer stay in school, time relates to money, and money relates to the utilization of human resources.

The point to be made here is that we must refrain from the rhetoric that casts suspicion on the motives of educators and policy makers who call for lengthened stays in school, longer days, increased instructional hours and the like. It is not a problem of democratic philosophy or a threat to equal opportunity that is at stake here. It is the very health of a democratic social system that should concern us.

2. The schools and economic policy. When all of the many reasons for an enlightened citizenry are examined, what clearly emerges is the central social function of the formal school. It is to give the individual the intellectual and vocational tools with which to survive, and to make of the social system, the recipient of the talent of its citizens. When the schools have been effective, both the society and the individual has prospered; when they have not, both the individual and the collective society live a leaner and less comfortable life. When the schools have been effective, people who have gainful

employment and are literate can enjoy a rich cultural environment. These are truisms, but are from time to time overlooked. It is little wonder, then, that in the opening paragraphs of A Nation At Risk we read:

Our Nation is at risk. Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world. ...

History is not kind to idlers. The time is long past when America's destiny was assured simply by an abundance of natural resources and inexhaustible human enthusiasm, and by our relative isolation from ... older civilizations. We live among determined, well-educated, and strongly motivated competitors. We compete with them for international standing and markets ... America's position in the world may once have been reasonably secure with only a few exceptionally well-trained men and women. It is no longer.

... Our concern, however, goes well beyond matters such as industry and commerce. It also includes the intellectual, moral, and spiritual strengths of our people which knit together the very fabric of our society. The people of the United States need to know that individuals in our society who do not possess the levels of skill, literacy, and training essential to this new era will be effectively disenfranchised, not simply from the material rewards that accompany competent performance, but also from the chance to participate fully in our national life. A high level of shared education is essential to a free, democratic society and to the fostering of a common culture, especially in a country that prides itself on pluralism and individual freedom. (pp. 5-7)

Policy makers at various levels of American life have argued that this country's economic future is in jeopardy because a large portion of our students are not learning the fundamental skills taught in the schools. This observation is compounded still further by charges from other quarters that not only are our students not learning the fundamental skills as effectively as they should be, but that if this country is to remain competitive world-wide, those skills which we have traditionally called, "fundamental" must now be redefined (Hunt, 1984).

There is another disturbing observation that should give one pause. The economic structure of the United States is undergoing a remarkable transition

from an industrial-based society, to one that is service-oriented. The service referred to by Hunt has more to do with information processing, data analysis, computer programming, and the like, than the more familiar forms of service related to food management, truck driving, and cleaning establishments. This movement in the predominant characteristic of the labor force is now calling forth such higher order thinking skills as analysis of data, the ability to select, organize and develop abstractions, and to deal with probabilities and the like.

Technology greatly affects how we do our work and other countries have become major competitors with the U.S. for the world market. The majority of workers in the United States are now engaged not in industrial production, but in service occupations such as trade, transportation, banking, public utilities, and so on. In 1950, 55% of the U.S. work force was employed by industry; today that figure has fallen to 20%. In contrast, in 1950, 30% of our work force held service and information jobs; today that figure has climbed to 70%.

...Certainly those jobs promising upward mobility will require the mastery of skills that go beyond today's definition of "the basics". To adequately prepare our young people for these jobs, we must expand our idea of the basics to include not only the abilities to read, write, and compute at a rudimentary level, but also the ability to put to creative use more complex skills. (Hunt, 1984, pp. 538-539)

But does a service-oriented economic system need these skills? Hunt, who is Chairman of the Task Force on Education for Economic Growth (a Task Force of the Education Commission of the States) admits that not all service jobs are, or will be, those which require higher-order thinking. However, he is quick to point to those jobs which promise the greatest opportunity for advancement, and these will require more complex skills, and the ability to apply them creatively in the marketplace.

Governor Hunt is not willing to cede the industrial production and advanced

technology to others. He issues a call for expanded economic production and enhanced technological capabilities. The inference, of course, is that a modern complex society such as this one is — with its past history as an industrial giant — cannot fall behind in a world that is becoming increasingly sophisticated. And this will mean extensive retraining of present workers and an increased emphasis on the development of a new set of "fundamental skills" that reflect a "robotics-oriented" industry, and a computer-centered commerce.

3. The pendulum effect. A serious problem affecting the development of educational policy in this country has always been the drastic shifts in educational "innovations" which periodically confront us. We are constantly confronted with the tension that arises from a desire for a stable educational policy on the one hand, and a concern that we not lose our ability to respond to needs, on the other. This has resulted in a crazy patchwork of methods and philosophies that emerge on the educational scene fairly rapidly, and hang on long after a new set of methods and philosophies emerge.

We are today drawn into a new set of discussions and plans; that of dealing with the need to develop tests of higher order skills. While I do believe that these discussions have been long over-due, I am confronted with the nagging doubt that we will reach an adequate resolution before we are off on still another path. Kaestle (1985) has addressed this problem in what I believe to be, a logical, rational way. He cautions against mistaking rhetoric for reality, and making premature plans and commitments.

Kaestle makes several points worth noting here:

1. An historical perspective of the swings in educational thought and behavior leads to the conclusion that these swings are inevitable. Our schools have accepted so many tasks that are inconsistent one with the other that it is a certainty that swings will occur between the inconsistent roles the schools have agreed to play.

2. The schools of this country when compared to those of others appear to be more effective in one period than in another, due to the shifting of themes and priorities in education.
3. The metaphor of the swinging pendulum is, to Kaestle, an inadequate one. He holds that the "reality of the system" is more stable than we think. He prefers to think not of a pendulum, but of a large, ungainly tanker that rocks a bit from side to side as it adjusts to currents.
4. While he tends to support a swing of the pendulum (or tanker) towards greater academic rigor, he returns to his main theme; the reality of American schooling is not well-served by the present rhetoric. He shows his ambiguity about American schooling by hoping that we not, in the process, abandon commitments to equality of opportunity, democratic access, and cultural reconciliation; as if a commitment to academic rigor would render these other commitments obsolete.

While I tend to admire the logic portrayed in Kaestle's paper, I cannot help but be struck by the uncertainty that affects him as it does all of us. Whether we call our shifts "pendulums" or we call them "freighters", the lack of stability of educational policies, and the inconsistencies of our political and social beliefs are preventing us from developing a well-thought through, carefully argued approach to what is best in schooling for all children. While we may never resolve all of the inconsistencies we carry with us, as with accumulated odds and ends stored in attics, some periodic culling is necessary.

Higher-order thinking. The nation's renewed interest in raising the level of expectations of students has brought to the fore, discussions of enhancing higher order reasoning skills (variously defined). So much confusion exists in these discussions that it is a wonder that cognitive psychologists, psychometricians, educationists, political leaders and the general public can understand each other at all.

In order not to contribute to this confusion any more than necessary, this paper will be limited to only a few facets of the present interest in

higher-order thinking. These topics were chosen because of their relationship to what is presently being proposed as "excellence" in education, and will include (1) the hierarchical nature of learning, (2) the impact of coaching on this hierarchy, and (3) some issues in the instruction of problem-solving techniques. The discussion which follows is constrained by lack of time and space and is meant to be only an indicator of the complexity of what is being widely discussed; it is by no means definitive.

1. An intellectual hierarchy. A major contribution to understanding the hierarchical nature of learning was made by Bloom and his colleagues, in a series of conferences held from 1949 to 1953 (Bloom, et al., 1956). These meetings resulted in a taxonomy of educational objectives dealing with the cognitive and affective domains (Bloom, 1956; Krathwohl, 1964). A third domain dealing with psychomotor skills was later addressed (Harrow, 1972).

More than thirty years after it was first conceived, the Taxonomy of Educational Objectives has once again emerged as a focal point in the planning of instructional programs and measurement of higher order skills and concepts. It is being used as it was intended; for communication and definition. The educational objectives listed in the taxonomy are organized into two large groupings; those dealing with the acquisition of knowledge, and those related to intellectual abilities and skills — currently referred to in some test manuals as "process" objectives. Within these two large groupings we find six major classes of objectives which are arranged hierarchically from relatively simple to complex; (1) Knowledge, (2) Comprehension, (3) Application, (4) Analysis, (5) Synthesis, and (6) Evaluation. Within each of these six classes are subclasses which are also arranged in order of complexity.

While not immediately related to this topic, those interested in addressing higher-order skills would do well to refer to the literature of cognitive

psychology which explores the hierarchical structure of general and special aptitudes (See for example, Vernon, 1960; Guilford, 1967; Humphreys, 1979). Unfortunately, more is known about the hierarchical structure of knowledge and abilities than is currently being used in the planning of instruction and measurement of higher order thinking.

2. The effect of coaching. Bloom, et al., admitted that while their hierarchy of educational objectives was a defensible one, it was by no means perfect. Of considerable concern to them was the role played by the student's prior experiences. They were troubled by the effect of prior training on the apparent position of an instructional objective in the hierarchy. For example, if a student was given an exercise designed to test a higher-order objective, and as a result of prior coaching resolved the exercise in an acceptable way, could it be assumed that the student was indeed demonstrating mastery of a higher-order objective, or simply exhibiting mnemonic recall (a skill rated lower in the hierarchy)?

Others have also turned their attention to the effect of coaching or other kinds of prior experiences on the nature of what is apparently being measured (Messick, 1980; Jensen, 1981; Doyle, 1983; Fredericksen, 1984; Anastasi, 1982). While Bloom and others dealt with the impact of coaching on the hierarchical position of an educational objective, some have dealt with the same issue from a different perspective. Jensen has been concerned with higher-order thinking as a facet of intelligence. He, along with others interested in cognitive psychology, has identified certain classes of problems which appear to require a higher-order of reasoning than do others, e.g., number series, verbal analogies, matrices items and the like. Using Spearman's connotation of a g factor (general intelligence), Jensen identifies these types of items as having a high degree of g loading. If a student who had no coaching on these item types were

to do well on them, chances are high that the student would also do well on a test of intelligence. If a student did not do well on these item types, initially, but was subsequently given coaching on them chances are that some modest improvement on item performance would be the result. However, that student would not do better on a test of intelligence even though some of these item types appeared in the intelligence measure. Jensen maintains that there is little transference of performance on a specific item type to another situation requiring a higher level of intellectual skill. The residual of that coaching would be recall, which is an intellectual skill of a lower order.

Is there a distinction between coaching, orientation to test-taking skills, and instruction in more general intellectual skills? Anastasi has discussed the difference between these three experiences (1982). She argues as does Jensen (1981), Messick (1980) and others that coaching may increase a test score slightly, but has no significant effect on the psychological or educational trait of interest. She observes that "...the more closely instruction is restricted to specific test content, the less likely is improvement to extend to criterion performance." (p. 42)

Anastasi views an orientation to test-taking procedures in a different light from that of coaching. She recognizes as have others, that a student who has had previous experience in taking tests has a certain advantage over those who have not had a similar experience. Booklets on how to take a particular kind of test, practice exercises in a test-like format which utilize none of the items that would appear on a test helps raise a student's "test-taking" level of sophistication but would have no effect on a student's memorization of a test-item response, and hence no measurable effect on the criterion being assessed.

This point of view is supported by recent research in the processing of

information (Schneider and Shiffrin, 1977; Schneider and Fisk, 1982; Shiffrin and Dumais, 1981). It has been shown that two visual-search tasks can be carried on simultaneously without affecting performance on a criterion of interest if one of the tasks can be performed automatically. If the procedure of test-taking is familiar to the student, then the skills connected with the mechanics of taking a test will not impact on that characteristic which is being measured.

And, finally, Anastasi draws a distinction between coaching and direct instruction in broad intellectual skills such as problem solving. Her support for this activity is a tentative one; she believes that not enough is yet known about whether this kind of training can result in an improvement of intellectual skills, as well as being able to transfer over to other content and settings.

3. Problem solving instruction. Fredericksen (1984) has recently completed an excellent review of the literature dealing with problem solving. His primary focus was on the theories and suggestions made by cognitive psychologists concerning how problem solving could be taught to others. The review covered research dealing with information processing, the nature of problems (those that were well defined as contrasted with those which were not), some aspects of problem solving theories, and studies of attempts to teach problem solving. A summary of Frederickson's conclusions after examining the research literature would indicate that:

1. It is possible, after considerable practice, to learn how to solve problems.
2. Problem solving involves mastery of appropriate and well-organized knowledge structures, a well-defined problem, automatic information processing, and the development of efficient pattern-recognition systems.
3. Problem solving is specific to a relatively narrow area of expertise; there appears to be little transfer from one

domain to another.

4. The skill of problem solving appears to be a sequential one proceeding from ill-structured problem recognition to well-structured problem identification; from slow, laborious information processing, to automatic processing. This sequential development is closely related to considerable practice which builds experience.
5. Generalizations about problem-solving techniques can only be made with some confidence when applied to well-defined problems met in such highly structured domains as mathematics, chess, and the like.
6. There is disagreement among those who write about problem solving processes concerning instructional methodologies to be employed. Some would argue that students should be taught problem solving methods through discovery; others call for explicit instruction.
7. While considerable agreement exists concerning the relationship of a student's experience to instructional methodologies employed, there is not this level of agreement when it comes to instruction of such special aptitudes as spatial, verbal, and reasoning ability.
8. It would appear from Fredericksen's review, that Anastasi's tentativeness concerning the transferability of general problem solving procedures will have to remain tentative. Fredericksen concludes that such highly specific areas as "electronics troubleshooting and auto repair" will require instruction in specific procedures because of the repetitious nature of the problem incurred. However, he assumes that more general instruction would be better for ill-defined and poorly structured problems that might be met in some future. (pp.391-398)

No discussion of the exercise of higher-order reasoning should begin and end with what specialists in cognitive psychology report; not because their work isn't important — it is very important. But since there is presently no wide-spread, effective link between psychologists and the schools, this problem needs to be addressed also, by those whose primary interests are education and schooling. How we think, has been a fundamental concern that stretches back into time and has been addressed by philosophers, psychologists and educators. While one could dwell on Aristotelian logic, and Socratic arguments, I would

accept a more modest sweep of history and call our attention to Dewey's, Democracy and Education (1916), and How We Think (1933). Much of what he addressed in these two volumes is still of interest to us seventy years later.

Two conclusions drawn by Dewey are fundamental to the question of thinking; (1) experience is a prime requisite to reflective thought, and (2) while experience is not primarily cognitive, the value of experience "lies in the perception of relationships or continuities which lead up to it" (Dewey, 1916; p. 164). From these two propositions flow a "method" of inquiry which is still the model used in research methodology; (1) a solid base of experience from which a problem is defined and addressed, (2) the development of alternative solutions or hypotheses related to the problem, (3) a collection of information needed to address the problem, (4) an analysis which leads to testing the alternative solutions or hypotheses and (5) conclusions drawn by the student which lead to the generation of "new" knowledge (Ibid, p. 192).

Dewey recognized the multi-faceted character of the thinking process, and was careful to address it in a comprehensive way. Those who write and discuss this problem today, tend to think of it too narrowly and dwell on it primarily from their small world of specialization. Practitioners run the risk of planning "solutions" (in the form of pedagogy and tests) from the perspectives of either psychologists, philosophers, politicians, or "gurus" whose eloquence may offer little more than inspiration or entertainment. What began as a desire to raise expectations of student achievement has moved into a national discussion of higher-order thinking processes that should be taught to all children in our schools. Somewhere along the line we may have lost our way.

This brief summary of some of the most current literature on higher order thinking processes should indicate that the greatest promise for the development of complex reasoning skills within students lies in the development of adequate

enough knowledge bases within a variety of subjects. It is apparent from past and on-going research that it is unrealistic to believe that the skills of analysis, prediction, application and evaluation can be practiced or taught outside of the content field. Knowledge of these content fields does not reside within books alone. Knowledge is experience, and whether that experience is attained through exercises in textbooks, television, in the laboratories of the schools, or outside the walls of the classroom, the ability to solve problems emerges from that experience and does not stand alone. If we are to broaden our perceptions of what constitutes an essential educational program for all of our students and a challenging one for our brightest, then our instruction and our measurement of that learning will reside primarily in the knowledge and experiences we offer them in the schools and the expectations we hold for their learning of that knowledge.

Educational measurement issues. At this point one might question why this paper has devoted so much space to a discussion of public policy and the characteristics of higher order thinking. The title of this paper would normally lead one to believe that what should be discussed are tests and how they can measure academic achievement at some level. As one who has spent the last thirty years writing standardized achievement tests, I believe that the greatest problem we face is not to be found in test construction, alone, but in test use.

A major context for contemporary test use has been the development of public policy and the monitoring of adherence to these policies. Given the preceding discussion of issues of egalitarianism and elitism, and the instability of social policy, it would seem to me that those who would use test results to demonstrate the need for excellence in education will find support for that position in the current literature. There is little doubt that the overall

record of academic performance exhibited by a large segment of American students is mediocre; particularly those presently enrolled in the secondary schools. While I agree with those who have argued that test scores taken by themselves are not a true and full indication of the health of schooling within a nation or a local school district (Lewis, 1985; Boyer, 1984), I do not agree that we cannot take a measure of school effectiveness using test data. If achievement test results are coupled with the use of input measures, such as a transiency index, some selected demographic data, and scholastic aptitude measures, we can legitimately draw conclusions about the status of schooling within a district or a building.

1. The nature of the test score decline. Test score declines have been a serious area of study for a decade or more (Harnischfeger and Wiley, 1975; Munday, 1979; Lerner, 1979; Tomlinson and Treacy, 1979; Wainer, 1984). Sederburg and Rudman (1984) analyzed the results of studies over the decade 1972-83. They found that the highest cohort of achievers showed the greatest declines across a variety of subjects as well as across age-level groups. There appeared to be a substantial drop among those children in the middle range of achievement, but less loss and some modest gains at the lower levels of achievement.

In one study reported by The Educational Testing Service, the national score average on the SAT's declined 10 points during the years 1976-1983. The decline for the top 10% of students was 26 points, while the lowest 40% of the test-taking population actually improved by 7 points. When these data were examined by academic subjects tested, the top 10% of the test-taking population showed a decline of 16 points in verbal skills, and 10 points in mathematics. The national average declined by 6 points in verbal skills, and 4 points in mathematics. The bottom 40% of the students taking the SAT showed gains of 4

points in verbal skills and 3 points in mathematics (Ramist, 1983).

In another study conducted, the percent of students who scored 600 or above on the SAT Verbal Skills subtest (moderately high achievers) fell from a high of 11.4% in 1972 to 6.8% in 1983. A similar drop in mathematics was noted from this group of achievers; from a high of 17.8% in 1972 to 14.4% in 1981. From 1981 to 1983 the proportion of students who scored 600 or above in mathematics increased from 14.4% to almost 16%.

A comparison was made of the results of two standardizations of the Iowa Test of Basic Skills, 1970, and 1977 (Munday, 1979). The relationship between gains and losses for the the three achievement groups under study remained consistent with other data examined, viz., that those who were in the high achieving group tended to show a greater decline in achievement than did those in the lower groups.

A recent report of the National Assessment of Educational Progress dealt with mathematics achievement in the three age groups traditionally studied; 9, 13, and 17 (NAEP, 1983). The four categories of mathematics skills examined were Knowledge, Skills, Understanding, and Applications. The first two categories are relatively low order skills, while the last two categories deal with abstractions and generalizations; higher-order intellectual skills. The results of the NAEP study were similar to other studies reported. It was clear that at the lower age levels, mathematics achievement had improved somewhat across both high and low ability cohorts. The percent of gain for the lower ability groups was more than double that of the gain for the higher ability cohort in Knowledge and Skills (lower level intellectual skills). The high ability group showed a small gain in Understanding, and no gain at all in Application skills. The low ability group showed a small though not significantly higher gain in Understanding as compared with the high group, and

showed a much higher percent of gain in Applications as contrasted with no gain shown by the high ability cohort.

During the years 1978-1982, those students in the high quartile of the 17 year-old age group scored lower in all of the mathematics categories in 1982 than they had four years earlier. The low quartile of 17 year-old students, in contrast, scored higher than they had four years earlier in the two mathematics categories that dealt with lower level skills and scored lower than they had four years earlier in the high order skills.

2. The impact of mandated assessment on achievement. The standardized achievement tests used in the schools for more than sixty years were not designed to drive a curriculum, but to reflect it (Rudman, 1977). And yet, since the late '60s and early '70s, mandated assessment tests have driven curriculums (Klein, 1984). The very nature of test construction has been heavily influenced by the pressure to use tests to monitor progress towards some predetermined goals; variously expressed as minimum cut-off scores, percentage of students attaining some proportion of mandated educational objective, a percentage of students attaining some score level deemed to be an indicator of excellence and the like (Rudman, 1985; Livingston and Zeiky, 1982; Jaeger, 1982).

There are those who would argue that stricter standards, more controls, and greater conformity will not motivate students and teachers to do better (Deci, 1985). And yet, the history of mandated assessment programs aimed at minimal performance, would indicate otherwise. These data would lead one to believe that the declines in test scores has been most rapid among the brightest students. I would infer from data such as these, that our previous assumption that high achieving students are inner-directed, self-motivating students who do not need as much instructional time as students who show lower school ability is

incorrect. Contrary to Deci's conclusion that stricter controls and higher standards will not motivate students to do better, it would appear that stricter controls on instructional time devoted to limited academic subjects taught to a specified target population (low achieving students) has brought about some improvement.

While it could be argued that all levels of education need enrichment, I fear that our nation is "at risk" when those students of high ability show a deterioration of academic achievement, for it is from this group that comes the future's intellectual and vocational leadership.

Another point that seems to emerge from test score analyses is the relationship between those subjects which have been tested by mandate, and the instructional time devoted to them. Approximately 80 percent of the states reporting state-wide concerns with minimal competencies are concerned with reading, and 70 percent with mathematics. Only 3 percent are seriously concerned with science, 18 percent with social studies, and 45% with language skills (Rudman, 1978). A correlate of this emphasis upon the testing of these subjects, is the amount of instructional time devoted by the schools. One national study of instructional time devoted to various subjects, indicates that in grades K-6, 89 minutes per day is devoted to the teaching of Reading, 50 minutes per day for Mathematics, 30 minutes for Social Studies, and only 24 minutes for Science (Weiss, 1978).

When teachers were asked to state their perceptions of their qualifications to teach Reading, Mathematics, Science and Social Studies, 63% felt very well qualified to teach Reading, 49% to teach Mathematics, 39% to teach Social Studies, and only 22% to teach Science. This variable, too, was an interesting correlate when compared to that which was selected as important to mandate. To what degree this feeling of instructional adequacy reflects true qualifications

and to what degree it is a function of familiarity is open to question (Weiss, 1978).

While there is strong indication that the high achieving student has shown a marked decline in achievement over time, it is equally clear that low achieving students have improved. It seems but a small jump in logic to conclude that the intensive concentration on the student needing remediation, and the attendant proportion of the day devoted to basic skill subjects has brought about an improvement in those subject areas and at those ability levels which have become, over the past twenty years, the targets for instructional focus. The major force in this improvement has been the political and public pressure on schools and their teaching staffs. We do have evidence that the schools can respond positively to this pressure. But we also have evidence that the nation has paid a price for this response in lowered expectations of students, simplified curriculums, inadequate instructional time devoted to useful subjects, and an imbalance in the human and fiscal resources used in schooling.

What Should We Do?

How then, can we address the pressures to raise our levels of academic excellence beyond the minimum? It seems to me that a convincing lesson learned in the past twenty years is that schools are not, by themselves, powerful enough to affect change on a national level. They have shown themselves to be effective "transmission belts" of public policy and are vulnerable to the pressures exerted by political leaders and such opinion-molders as the press. They are, in the words of an eminent sociologist, "maintenance" institutions (Parsons, 1958). If schools are not proactive, we need to assure the conditions which will allow them to be reactive. One powerful tool for accomplishing this has been state-wide assessment programs. If they have had success in focusing instruction for those needing remediation, they most assuredly will work as well

for identifying those students who need some other level of instruction. What follows is based upon certain assumptions. If these assumptions are reasonable, in light of what we seem to know, then the suggestions which flow from them are also reasonable.

1. The power of state-mandated assessment programs has been convincingly demonstrated to be a force in instituting instructional change within the schools.
2. Higher order reasoning skills can best be taught as an integral part of some specified body of knowledge.
3. There is a demonstrable relationship between focused instruction and student performance at all levels of ability.
4. Schools can be effective if the mandate given them is strong enough and if adequate resources are available.
5. There is a limit to the amount of resources — human and fiscal — that are available for education.
6. The schools are an important instrument in affecting social and economic policies of a nation.
7. Recommendations for reform of any institution, including the schools, must be based on a reasonable expectation of stability of public policy.

While the National Commission on Excellence in Education's report has had a significant impact on the public dialogue about educational standards in the schools, like other "public events" the conditions described may persist long after the public's attention has drifted elsewhere. The challenge we have before us is to keep the public's support and interest riveted on A Nation At Risk by responding to it in an aggressive and meaningful way. The following suggestions focus on the State of Michigan's educational program, but some of these ideas are applicable to other states as well.

The state should develop a plan which incorporates a two-tier evaluation of student academic status. While it is clear that mandated assessment programs

have focused on minimal achievement and have had some impact on lowering expectations of all students, this thrust should not be discontinued since test data would indicate that there still are too many students unable to meet defined minimal standards. Present plans in Michigan call for an expansion of the minimal competency test to areas other than Reading and Mathematics. I believe this to be an unwise move.

A modification should be made in the present Michigan Educational Assessment Program (MEAP); I would suggest that it be mandated on a matrix sampling basis rather than on an every-pupil requirement. The test results have been widely used by teachers and administrators for examining the "effectiveness" of instruction in Reading and in Mathematics. However, close scrutiny of the reliability indices of the items which comprise each objective would reveal that while they are remarkably high, given the few items per objective match, they are by no means high enough to warrant serious instructional planning for individual students. They are reasonably high enough, however, for planning of instructional programs for groups.

Matrix sampling could yield useful information for public monitoring of minimal achievement within the state's schools while at the same time reducing the amount of testing time, and expenditures connected with testing. Test results could be available on a state-wide, district, and building level (in those instances where the grade population within a building was adequate for this type of sampling). I would suggest that matrix sampling of Reading and Mathematics be given in grades 4, 7, and 9. This would permit an adequate sampling of the primary, upper elementary and middle school years. The contents of this minimal competency test should be held to the first two levels of the Bloom taxonomy (Knowledge and Comprehension). It seems clear from what we know that more abstract levels of thinking depend upon more than minimal knowledge.

A second tier of testing should be undertaken by the Michigan Department of Education. This testing should be mandated on an every-pupil basis in grades 1,3,5,8,10, and 11. The content of these tests should include much more than Reading and Mathematics. It should measure the language skills, social science, science, and listening skills at the appropriate grade levels. The content of these tests should consist of levels 1 to 4+ of the Bloom taxonomy (depending upon the grade level at which the test is administered). It should be so constructed that there are sufficient items at a variety of difficulty levels from 30-90. These tests should contain subtests with reliability coefficients in the .80's to .90's, with adequate item discriminations (.40+), and demonstrate content validity based upon the prevailing instructional materials in use in the schools. Approximately 50% of the test should consist of levels 1 to 3 of the Bloom taxonomy. In short, the quality of tests that have been available and used nationally for more than 60 years.

This second tier of tests could be used for the identification of different levels of scholastic ability which, in turn, would permit (1) grouping students for differential instruction, (2) determining the pacing of instruction in heterogeneous classrooms, (3) reporting reliable information to parents during conferences, (4) curriculum evaluation and planning, and (5) reliable disaggregation of subtests into the content clusters of which they are comprised. These tests could be used to report student achievement on an objectives-referenced basis, and with the help of a standard-setting mechanism to establish cut-scores for those levels of achievement deemed, "essential", and for those levels deemed as necessary for high achieving students.

The state should establish a standard-setting advisory committee. This committee would be comprised of teachers at the grade levels at which the test would be used, subject matter specialists familiar with the performance of

students on specific subjects at particular grade levels, and measurement specialists familiar with the methodology of standard setting. This committee's sole purpose would be advising the state's Department of Education on appropriate methods of establishing standards using test scores, and would determine two levels of standards; one would be related to that "essential" level of academic achievement that most students should attain at the various grade levels tested, and that level of "distinction" to be attained by some proportion of high achieving students.

Standards of academic performance have traditionally remained the prerogative of local school districts. I would hold that this should remain a responsibility of the local school district. But I would suggest a modification of this tradition. High school diplomas should be presented to those students who have successfully completed the local school district's program. But since there appears to be some divergence of standards among the local districts, I would suggest an endorsement by the state attached to the diplomas of those students who had successfully demonstrated an "essential" level of competence as determined by some passing score on a state-administered test. The level to be chosen would depend upon the judgments of the state-wide advisory committee on standard-setting.

I would suggest that the second tier of tests mentioned in the preceding paragraphs, be the vehicle for supplying this information. All students would take tests in grade 10. If they did not reach an "essential" level in grade 10, they would have an opportunity to retake portions of the test in the succeeding two years. If they failed to achieve this level, they could, if the school district so desired still receive a diploma, but one without the state endorsement.

The state should assume a major portion of the funding for mandated

assessments. If we can agree on the positive effect of educational change brought about by a mandated assessment program, then a dramatic increase in state funding should be made for using this tool for change. We cannot assume an unlimited supply of funds for any state-sponsored activity, so we obviously face some difficult reallocations of existing or even expanded resources. I am suggesting that the state adopt a two-tier system of assessment for grades 1-10, and in some instances, 1-12. This recommendation is not made lightly, nor without an understanding of its ramifications. The willingness of the state to take on this responsibility of full-funding will be a measure of its serious intent to support a commitment to excellence in education. The two-tier approach recommended in this paper still leaves some holes in the testing program; grades 2,4,6,7, and 9 will not have comprehensive information available to them. Local districts may want to supplement their state-funded program by allocating some local funding for these grades.

A major proportion of operational funds allocated for education are committed to salaries for professional personnel. In one study of state-wide expenditures for instructional materials and testing (Rudman, 1976), approximately .5 of 1% was spent on the acquisition of new instructional materials, and .03 of 1% was used to purchase test booklets, scoring services and answer sheets. Approximately 80% was spent on salaries. While I do not quarrel with adequate salaries for professionals, I do feel that not nearly enough is being spent on those components of an instructional program that students and teachers need. What is required is some courageous reallocation of funds, and a new look at priorities for educational expenditures.

The state Department of Education should establish a technical advisory committee. This committee would serve a variety of functions related to the determination of test specifications, criteria for selection of tests, and

recommendations for choosing tests or test development contractors. Test development issues could be explored in detail by this committee and aired before decisions are made which would prevent alternatives to state developed or state-sponsored development of tests.

There is a trend developing within the United States for state and local school districts to develop their "own" tests. These tailor-made tests are assumed to have higher content validity, because they most closely relate to the instructional goals set by the users of these tests. There has also been a notion that the application of item response theory to test development has increased the possibility that tests can be constructed by "stitching together" items from various sources or item banks and by relating these items to the measurement of specific educational objectives. While item response theory has some limited applications in the scaling of items, it is not widely accepted as an appropriate use for test construction (Traub, 1983; Hoover, 1984). The result, too often, is a collection of items which are mistakenly called, a test.

The technical advisory committee could help prevent errors too often made by state departments of education. All too frequently, some of these departments have thrust ahead on test development without appropriate concern for the relationship of proper test development procedures. It would seem as if each state is convinced that the subject matter taught in its schools is unique; ignoring the fact that much of the subject matter content taught in the schools has a common sequence of development. This is reflected in the instructional materials used in the schools and should be related to specific test content.

The technical advisory committee would be the appropriate group to examine the educational and psychometric specifications of tests chosen or developed by the state's Department of Education. This would include recommended levels of difficulty and discrimination, adequate sampling of subject matter domains, and

an appropriate proportion of the tests content devoted to levels 3 and above of the taxonomy of educational objectives.

Summary

It has taken awhile for the full import of A Nation At Risk to be realized. Citizens of this nation from all walks of life, from all educational levels are troubled by what they have seen or experienced in the schools. Not all of the criticisms are valid, not all are rational. But the evidence is too compelling to ignore. We do seriously need to reform our expectations of what education can do for the welfare of the individual and the welfare of the nation.

There is a countervailing movement afoot to ridicule and discredit these concerns. The point has been made and documented that school reform is necessary and that our expectations must be raised. This does not mean, contrary to what we are just beginning to hear, a movement to elitism in our schools or an abandonment of children who live impoverished lives. We have spent the last twenty years intensively addressing the needs of these children. These twenty years have covered the life spans of those who are presently touched by institutions of learning. We have shown that to some modest degree the attention we have lavished on minimal competency programs has resulted in improved achievement for this cohort. But we have also shown that this has not been without some cost.

The recommendations made in the last section of this paper are few in number and not impractical. Whether we can effectively address the educational needs of all children has yet to be determined. But of one thing I am certain; if more attention is not paid to those students who need enriched curriculums, and challenging standards to meet, every person in this nation will be diminished in terms of accessing the material and cultural rewards that living should bring.

REFERENCES

- Admissions Testing Program of the College Entrance Examination Board. National Report on College-Bound Seniors, 1975-76. (Princeton: Educational Testing Service, 1975).
- Admissions Testing Program of the College Entrance Examination Board. National Report on College-Bound Seniors, 1976-77. (Princeton: Educational Testing Service, 1976).
- Anastasi, Anne. Psychological Testing Fifth Edition. (New York: Macmillan Publishing Company, Inc., 1982).
- Beyer, Barry K. "Improving Thinking Skills — Defining the Problem", Phi Delta Kappan. 65:486-490, (March, 1984).
- Beyer, Barry K. "Improving Thinking Skills — Practical Approaches", Phi Delta Kappan. 65:556-560, (April, 1984).
- Bloom, Benjamin S. (Editor). Taxonomy of Educational Objectives Handbook I: Cognitive Domain. (New York: David McKay Company, Inc., 1956).
- Boyer, Ernest L. "Reflections on the Great Debate of '83", Phi Delta Kappan. 65:525-530, (April, 1984).
- Brandenburg, Richard K., James J. Gallagher, and Glenn D. Berkheimer. Science and Mathematics Preparation for Post-Secondary Education: The Implications of Declining Test Scores. Bulletin No. 1-78, (East Lansing, Michigan: Science and Mathematics Teaching Center, Michigan State University, 1978).
- Clark, Burton R. "The High School and the University: What Went Wrong in America, Part I", Phi Delta Kappan. 66:391-397, (February, 1985).
- de Bono, Edward. "The Direct Teaching of Thinking As A Skill", Phi Delta Kappan. 64:703-708, (June, 1983).
- Deci, Edward L. "The Well-Tempered Classroom", Psychology Today. 19:52-53, (March, 1985).
- Dewey, John. Democracy and Education. (New York: The Macmillan Company, 1916).
- Dewey, John. How We Think. (Boston: D.C. Heath, 1933).
- Doyle, Walter. "Academic Work", Review of Educational Research. 53:159-199 (Summer, 1983).
- Duncan, James K. "Traditions in Educational Evaluation — American Style", The Educational Forum. 44:137-145 (January, 1980).
- Eberly, Charles. "Orientation Test Scores of Fall, 1976, New Freshman (Fall, 1976 New Freshman Compared to New Freshman in Recent Years)" mimeographed, 1976.

- Fredericksen, Norman. "Implications of Cognitive Theory for Instruction in Problem Solving", Review of Educational Research. 54:363-407 (Fall, 1984).
- Gipps, C. and Harvey Goldstein. Monitoring Children: An Evaluation of the Assessment of Performance Unit (London: Heinemann, 1983).
- Goldstein, Harvey. "Measuring Changes in Educational Attainment Over Time: Problems and Possibilities" Journal of Educational Measurement 20:369-377 (Winter, 1983).
- Governor's Educational Summit Task Force. (Utica, Michigan: The Task Force, 1985).
- Guilford, J.P. The Nature of Human Intelligence. (New York: McGraw-Hill, 1967).
- Harnischfeger, Annegret and David Wiley. Achievement Test Score Decline: Do We Need to Worry? (Chicago: ML — Group for Policy Studies in Education, CEMREL, Inc., 1975).
- Harrow, Anita J. Taxonomy of the Psychomotor Domain. (New York: David McKay Company, Inc., 1972).
- Hoover, H.D. "The Most Appropriate Scores for Measuring Educational Development in the Elementary Schools: GE's", Educational Measurement. 3:8-14 (Winter, 1984).
- Humphreys, L.G. "The Construct of General Intelligence", Intelligence. 3:105-120, (1979).
- Hunt, James B. "Education for Economic Growth: A Critical Investment", Phi Delta Kappan. 65:538-541 (April, 1984).
- Jaeger, Richard M. "An Iterative Structured Judgment Process for Establishing Standards on Competency Tests: Theory and Application", Educational Evaluation and Policy Analysis. 4:461-475 (Winter, 1982)
- Jensen, Arthur R. Straight Talk About Mental Tests. (New York: The Free Press, 1981).
- Kaestle, Carl. "Educational Reform and the Swinging Pendulum", Phi Delta Kappan. 66:422-423 (February, 1985).
- Klein, Karen. "Minimum Competency Testing: Shaping and Reflecting Curricula", Phi Delta Kappan 65:565-567 (April, 1984).
- Krathwohl, David. (Editor). A Taxonomy of Educational Objectives. Handbook II: Affective Domain. (New York: David McKay Company, Inc., 1964).
- Lerner, Barbara. "Tests and Standards Today: Attacks, Counterattacks, and Responses.", in Roger T. Lennon (Editor) New Directions for Testing and Measurement. (San Francisco: Jossey-Bass, Inc., 1979).
- Lewis, Anne C. "Test Scores, Test Scores on the Wall ..." Phi Delta Kappan 66:387-388 (February, 1985).
- Lipman, Matthew. "Philosophy for Children". National Forum. 65:18-21 (Winter,

1985).

Livingston, Samuel A. and Michael J. Zieky. Passing Scores. (Princeton: Educational Testing Services, 1982).

Messick, Samuel. The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research From the Fifties to the FTC. (Princeton: Educational Testing Services, 1980).

Michigan Commission on High Schools. Preliminary Report. (Lansing, Michigan: The Michigan Association of Secondary School Principals and the Michigan Department of Education, 1983).

Michigan Republican Party Central Committee. Excellence in Education: A Republican Action Plan for the '80s. (Lansing, Michigan: The Committee, 1983).

Michigan State Board of Education. Better Education for Michigan Citizens: A Blueprint for Action. (Lansing, Michigan: State Board of Education, 1984).

Morante, Edward A. and Anita Ulesky. "Assessment of Reasoning Abilities", Educational Leadership 42:71-74 (September, 1984).

Munday, Leo A. "Changing Test Scores, Especially Since 1970", Phi Delta Kappan 60:496-499 (March, 1979).

Munday, Leo A. "Changing Test Scores: Basic Skills Development in 1977 Compared with 1970", Phi Delta Kappan 60:670-671 (May, 1979).

National Assessment of Educational Progress. The Third National Mathematics Assessment: Results, Trends and Issues. (Denver, Colorado: Education Commission of the States, 1983).

National Commission on Excellence in Education. A Nation At Risk (Washington, D.C.: U.S. Government Printing Office, 1983).

Paul, Richard W. "The Critical-Thinking Movement", National Forum 65:2-3,32 (Winter, 1985).

Parsons, Talcott. "Suggestions for a Sociological Approach to the Theory of Organizations—I", Administrative Science Quarterly 1:63-85 (June, 1956).

Parsons, Talcott. "Suggestions for a Sociological Approach to the Theory of Organizations—II", Administrative Science Quarterly 1:225-223 (September, 1956).

Pipho, Chris. "The Momentum of Education Reform: States on the Move", Phi Delta Kappan 66:85-86 (October, 1984).

Ramist, Leonard and Solomon Arbeiter. Profiles, College-Bound Seniors, 1982. (New York: College Entrance Examination Board, 1983).

Ramist, Leonard and Solomon Arbeiter. Profiles, College-Bound Seniors, 1983. (New York: College Entrance Examination Board, 1984).

- Rudman, Herbert C. "The Superintendent and Testing: Implications for the Curriculum", Measurement in Education 8:1-8 (Fall, 1977).
- Rudman, Herbert C. "The Standardized Test Flap", Phi Delta Kappan 59:179-185 (November, 1977).
- Rudman, Herbert C. Stanford Achievement Test: Organizational Planning Guide. (Cleveland, Ohio: The Psychological Corporation, 1985).
- Rudman, Herbert C. "Science Teaching in the '80s — A Case of Benign Neglect?" Science and Children. 16:7-8 (October, 1978).
- Schneider, W. and A.D. Fisk. "Concurrent Automatic and Controlled Visual Search: Can Processing Occur Without Resource Cost?", Journal of Experimental Psychology: Learning, Memory and Cognition 8:262-278 (1982).
- Schneider, W. and R.M. Shiffrin. "Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention." Psychological Review. 84:1-66 (1977).
- Scriven, Michael "Critical for Survival", National Forum 65:9-12 (Winter, 1985).
- Sederburg, William A. and Herbert C. Rudman. "Educational Reform and Declining Test Scores", Michigan School Board Journal 38:8-10,24 (April, 1984).
- Shane, Harold G. "The Academic Score Decline: Are Facts the Enemy of Truth?" Phi Delta Kappan 59:83-86, 145-146 (October, 1977).
- Shiffrin, R.M. and S.T. Dumais. "The Development of Automatism" in J.R. Anderson (Editor). Cognitive Skills and Their Acquisition. (Hillsdale, New Jersey: Erlbaum, 1981).
- Start, K.B. and B.K. Wells. The Trend of Reading Standards (Slough: The National Foundation for Educational Research, 1972).
- Tomlinson, Tommy M. and Maureen Treacy (Editors). Achievement Testing and Basic Skills: Conference Proceedings. (Washington, D.C.: U.S. Government Printing Office, 1979).
- Traub, Ross E. "A priori Considerations in Choosing an Item Response Model", in Ronald K. Hambleton (Editor). Applications of Item Response Theory. (Vancouver, B.C.: Educational Research Institute of British Columbia, 1983).
- Vernon, P.E. The Structure of Human Abilities. (London: The Macmillan Company, 1960).
- Wainer, Howard. "An Exploratory Analysis of Performance on the SAT", Journal of Educational Measurement 21:81-91 (Summer, 1984).
- Weiss, Iris R. 1977 National Survey of Science, Mathematics, and Social Studies Education. (Research Triangle Park, North Carolina: Center For Educational Research and Evaluation, 1978).

Wirtz, Willard (Chairman). On Further Examination. (New York: College Entrance Examination Board, 1977).